AD_____

Award Number: W81XWH-13-1-0064

TITLE: A Structural Biology and Protein Engineering Approach to the Development of Antidotes against the Inhibition of Human Acetylcholinesterase by OP-based Nerve Agents

PRINCIPAL INVESTIGATOR: Joel L. Sussman

CONTRACTING ORGANIZATION: Weizmann Institute of Science
Rehovot 76100 ISRAEL

REPORT DATE: March 2014

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
March 2014	Final	11 February 2013 - 10 February 2014

4. TITLE AND SUBTITLE
A Structural Biology and Protein Engineering Approach to the Development of Antidotes against the Inhibition of Human Acetylcholinesterase by OP-based Nerve Agents

5a. CONTRACT NUMBER
W81XWH-13-1-0064

5b. GRANT NUMBER
W81XWH-13-1-0064

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Joel L. Sussman – PI
Dan S. Tawfik – Co PI

E-Mail: Joel.Sussman@weizmann.ac.il

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Weizmann Institute of Science
Rehovot 76100 ISRAEL

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The human acetylcholinesterase (hAChE) gene was cloned into the pHLsec expression vector. The recombinant enzyme (rhAChEmD) was expressed on a large scale in adherent 293T cells. It was secreted as a monomeric species, purified by affinity chromatography, and deglycosylated with PNGase F. A crystallization screen, using the Mosquito crystallization robot, identified conditions for formation of diffraction-quality crystals, in 0.025% dichloromethane/12% PEG 20,000/0.1M imidazole, pH7.0. The crystals formed in the hexagonal space group $P3_112$, with cell constants a=125.31, b=125.31, c=131.40 Å, and one monomer per asymmetric unit. A complete dataset, to 2.9 Å resolution, was collected at 100 K at the ESRF (Grenoble, France). The structure was solved by molecular replacement, resulting in an R_{free} of 23.74%, and R_{work} of 19.31%, for all data to 2.9 Å. The coordinates and structures factors have been deposited in the PDB, with IDcode 4PQE. As an initial step for synthesis of a GF surrogate, the chloridate, $CH_3P(O)$(O-cyclohexyl)Cl, has been prepared, and will be reacted with coumarin. Conditions have also been developed for expression of a full-length rhAChE construct, $rhAChE_T$ that assembles to form the physiological tetramer.

15. SUBJECT TERMS
human acetylcholinesterase; cloning; expression; X-ray structure; 3D structure; chemical synthesis; GF surrogate

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	22	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

1. INTRODUCTION

Human acetylcholinesterase (hAChE) is the primary target of organophosphorus nerve agents. The major benefit of the structural information provided by this research is to yield one or more products based on a combination of hAChE with one or more oxime reactivators that, together, will be capable of pseudocatalytic hydrolysis of all V- or G-agents. It is envisaged that these products will be utilized either prophylactically or in a post-treatment context. Embedded in a suitable medical doctrine they will expedite the recovery and return to full battle capability of the war-fighter. The approach adopted should substantially reduce the amount of the enzyme protein required to afford protection, thus making its availability more economically feasible. In parallel, the civilian population may be exposed to OP nerve agents both in military and terrorist scenarios. However, civilians are exposed to and susceptible to a variety of commercially available OP-based and carbamate-based insecticides quite routinely. The same considerations outlined above may also apply to treatment of civilians. The issue of reduced protein dosage, resulting in enhanced cost effectiveness, will be particularly relevant to the civilian arena, where treatment will be administered to a heterogeneous population in terms of age and medical background.

2. BODY

Overexpression of rhAChE in human embryonic kidney cells, H293T

The hAChE gene was cloned into the pHLsec expression vector. One sequence, encoding residues 1-574, results in expression of a catalytic subunit monomer (Velan et al, 1991). The full-length sequence, residues 1-614, contains at its C-terminus residues involved in subunit association for which two constructs were generated. The hAChEmD construct includes the vector secretion signal, followed by the sequence of the mature protein starting at Glu32, while the hAChE_M construct includes the native secretion signal followed by a 6His-Gly-TEV protease cleavage site, and then by the mature protein starting at Glu32. The rhAChEG4 construct is similar to hAChE_M, except that it contains the full-length catalytic polypeptide chain (Fig.1). Expression of the 3 constructs was tested on a small scale (10 cm plates). Adherent 293T cells were transfected using PEI as a transfection reagent. Transiently expressed protein in the medium was examined 3 and 5 days post-transfection. An example of such a small-scale expression test is shown in Fig. 2. Large-scale production utilized 64 large plates (15 cm) of adherent 293T cells. Transfected cells were grown in serum-free medium, and medium containing the secreted protein was collected 3 and 6 days post transfection. Following centrifugation and filtration, the media were taken for purification of the rhAChE (see below). Yields of protein are in the range of 3-7 mg per large-scale batch. rhAChE_M was produced 4 times, rhAChEmD was produced 8 times, with one prep yielding crystals that permitted structure determination (see below), and rhAChEG4 was produced 10 times, since the yields were lower for this construct.

Purification and deglycosylation of rhAChEmD

rhAChEmD was purified by affinity chromatography, and deglycosylated with PNGase F, so as to improve the chances of obtaining crystals diffracting to high resolution.

Affinity resin: The principal purification step was affinity chromatography on a Sepharose column to which a reversible AChE inhibitor had been attached via a suitable spacer (Sussman et al, 1988), For this purpose, *m*-[ε-aminocaproyl-ε-aminocaproyl)-aminophenyltrimethylammonium] bromide (mPTA) was coupled to CNBr-activated Sepharose 4B.

Purification of PNGase F: PNGase F removes N-linked polysaccharide chains from the asparagine residues to which they are attached in glycoproteins. We utilized recombinant *Flavobacterium meningosepticum* PNGase F (Grueninger-Leitch et al, 1996). Since high concentrations of the glycosidase are required for efficient action on native protein substrates, we chose to express and

purify it 'in house'. It was over-expressed as a glutathione S-transferase fusion protein – GST PNGase F - in *E coli*. The crude extract was passed over a Glutathione Sepharose High Performance (GSTrap HP) column from which the bound GST-PNGase F was eluted with a gradient of reduced glutathione. The pooled peak fractions were further purified by size-exclusion chromatography on a Sephacryl S-300 column.

Purification of rhAChE: The rhAChEmD construct used was a truncated one that produced monomers, and was devoid of tags (see above). The following is a typical purification:
The initial step was a batch procedure. 2.5 l of supernatant from the transfected HEK293 cells were incubated with 20 ml of the *m*PTA-Sepharose 4B affinity resin overnight at 4°C in a cell spinner. The following day, the Sepharose beads were separated from the medium by filtration on a sinter glass, and packed into a column. The filtrate was passed over the packed column, which was then washed with 400 ml of a 3-fold dilution of Buffer A (100 mM NaCl/10 mM Tris, pH 8.0), followed by 70 ml of undiluted Buffer A. The bound enzyme was then eluted with 50 mM tetramethylammonium bromide (TMA) in Buffer A. Fractions of 1ml were collected, and the protein peak was detected by monitoring at 280 nm. Based on assay of enzymatic activity using the Ellman procedure, with acetylthiocholine (ATC), as substrate (Ellman et al, 1961), ~2.3 mg of rhAChE were collected (assuming a specific activity of 6000 units/mg). Activity in a 15 ml of the purified enzyme amounted to 70% of the total activity in the crude supernatant. The sample was then concentrated to 4 ml in an Amicon Ultra-4 Centrifugal Filter Unit equipped with an Ultracel-30 membrane. Figs. 3 and 4 show, respectively, non-denaturing and SDS-PAGE gels for the purified rhAChE.

Deglycosylation of rhAChE: The entire concentrated batch (in 50 mM TMA/Buffer A) was deglycosylated as follows: An aliquot of 40 µl of 0.5mg/ml GST-PNGase F was added (rhAChEmD/PNGase protein ratio ~1:100). After incubation for 70 h at 26°C, the GST-PNGase F was removed by addition of 200 µl of Glutathione Sepharose 4 Fast Flow. After incubation at room temperature for 2 h, the beads were spun down at 3,000 g for 10 min. The supernatant was dialyzed overnight against 100 mM NaCl/1 mM MES, pH 6.5. SDS-PAGE revealed a significant decrease in the molecular weight of the polypeptide, as well as a decrease in the intensity of glycostaining (Fig. 5). MS revealed a decrease in mass of 1.1 kD, corresponding to the removal of 7-8 sugar residues.
Concentration of the deglycosylated rhAChEmD was performed in two steps:

1. It was concentrated to 1 ml in an Amicon Ultra-4 Centrifugal Filter Unit equipped with an Ultracel-30 membrane.
2. The 1ml concentrate was further concentrated in a Vivaspin 500 Centrifugal Concentrator, equipped with a MWCO 30 kD filter, to a final volume of 200 µl, and final protein concentration of 8.2 mg/ml.

Purification, tetramerization and deglycosylation of rhAChE$_T$

Affinity purification of rhAChE$_T$: Full-length rhAChE utilized the following construct: Signal Peptide(1-31)-His$_6$-TEV-AChE(32-614). The hAChE sequence is that of the hAChE$_T$ splice variant, which bears the WAT tetramerization sequence at its COOH-terminus. TEV is a cleavage motif recognized by the TEV protease, which permits removal of both the signal peptide and the His-tag.
Purification of the rhAChE$_T$, which had been expressed in HEK 293 cells, utilized affinity chromatography on the *m*PTA-Sepharose 4B affinity resin as the principal step. It was followed by deglycosylation in order to improve the chances of obtaining crystals diffracting to high resolution. The overall procedure was performed three times, and one such procedure is described in detail below.

The initial purification step was a batch procedure. A volume of 5-10 l of supernatant from the transfected HEK293 cells was incubated with 20 ml of the affinity resin overnight at 4°C in a cell spinner. The Sepharose beads were then separated from the medium by filtration on a sinter glass, and packed into a column. The filtrate was passed over the column, which was then washed with 400 ml of a 3-fold dilution of buffer A (100 mM NaCl/10 mM Tris, pH 8.0), followed by 70 ml of undiluted

buffer A. The bound enzyme was then eluted with 2 mM decamethonium bromide. The peak fractions of the eluate were pooled, and concentrated to 3.5 ml in an Amicon Ultra-4 Centrifugal Filter Unit equipped with an Ultracel-30 membrane.

Characterization of rhAChE$_T$: The purity and homogeneity of the affinity-purified rhAChE$_T$ were assessed by SDS-PAGE and native PAGE, by size exclusion chromatography (SEC), and by sucrose gradient centrifugation. Fig. 5 shows SDS-PAGE of the purified rhAChE$_T$, alongside *Torpedo californica* AChE (*Tc*AChE) and human butyrylcholinesterase (hBChE), under non-reducing and reducing conditions. Under non-reducing conditions part of the rhAChE$_T$ migrates as a dimer, and part as a monomer (Fig. 6). Under reducing conditions only the monomer is seen. Fig. 7 shows PAGE of the purified rhAChE$_T$ under non-denaturing conditions, again alongside *Tc*AChE and hBChE. It clearly migrates as a single band. The lower panel in Fig. 8 shows size exclusion chromatography (SEC) of the purified rhAChE$_T$ on a Superdex column. For comparison, the lower panel shows the pattern obtained with the purified G_2 *Tc*AChE dimer. The major peak of the rhAChE$_T$ elutes at the same column volume as the *Torpedo* dimer. However, a substantial peak elutes earlier, which most likely corresponds to the G_4 tetramer. Sucrose gradient centrifugation, performed on preparations different from that described above, revealed, for one batch 5.2 and 10.0 S peaks, and, for another batch, 4.9 and 9.9 S peaks. The heavier peak corresponds to the G_4 tetramer, while the lighter one may correspond to either monomer or dimer.

New rhAChE constructs: New constructs are being generated in order to obtain higher yields of purified protein. The hAChE_M is being modified to contain a 3-amino-acid linker between the His tag and the TEV site. We encountered difficulties in achieving efficient cleavage of the His-tag by the TEV protease in the current construct. We suspect that the inefficient cleavage is due to their being only one amino acid (Gly) between the His tag and the TEV site in this construct. The modified construct of the hAChE monomer should permit higher yields. The same modified construct will be cloned into a second vector that is compatible with transfection of 293E cells. 293E cells are grown in suspension; thus, handling larger volumes of culture for production will be easier.

Deglycosylation of rhAChE$_T$G$_4$: Deglycosylation with PNGase F was performed similarly to deglycosylation of the monomer, as follows. The GST-PNGase F was added to 2 mg of rhAChE$_T$G$_4$ in 3.5 ml of buffer A, at a PNGase F/hAChE protein ratio of ~1:100. Deglycosylation was performed for 72 h at 26°C. The GST-PNGase F was removed by addition of 200 µl of Glutathione Sepharose 4 Fast Flow beads. After incubation at room temperature for 2 h, the beads were removed by centrifugation at 3,000 g for 10 min. The supernatant was dialyzed overnight against 100 mM NaCl/1 mM MES, pH 6.5, and concentrated to 1 ml in an Amicon Ultra-4 Centrifugal Filter Unit equipped with an Ultracel-30 membrane. It was further concentrated in a Vivaspin 500 Centrifugal Concentrator, equipped with a MWCO 30 kD filter, to a final volume of 32 µl, at a concentration of 26 mg/ml of DG-rhAChE$_T$G$_4$.

Tetramerization of DG-rhAChE$_T$: The proline-rich attachment domain (PRAD) is the sequence within ColQ (the collagen tail that anchors AChE to the basal lamina within the synaptic cleft) around which the catalytic subunits of the enzyme assemble into the physiological tetramers via their COOH-terminal T-peptide (WAT) sequences (Fig. 9) (Dvir et al, 2004). Indeed, addition of synthetic PRAD to preparations containing AChE monomers and dimers *in vitro* promotes tetramerization in the test tube (Giles et al, 1998; Kronman et al, 2000).

In order to generate homogeneous G_4 rhAChE$_T$ for crystallization trials, PRAD was added in a slightly greater than 10-fold molar excess, *viz.,* at a final concentration of $1x10^{-4}$ M, to $0.86x10^{-5}$ M DG-rhAChE$_T$ (0.56 mg/ml). Prior to addition of PRAD, sucrose gradient centrifugation revealed two peaks, at 5.2 and 10.0 S, a pattern very similar to that observed prior to deglycosylation. However, in the presence of PRAD, the lighter peak was absent, and only a single 10.1 S species was seen (Fig. 10).

Synthesis of OP Nerve Agent Surrogates

As an initial step for synthesis of a GF surrogate, the chloridate, $CH_3P(O)$(O-cyclohexyl)Cl, has been prepared, and will be reacted with the corresponding coumarin.

Crystal Structure Determination

The concentrated solution of rhAChEmD was subjected to a series of crystallization screens, via the sitting drop method, using the Mosquito crystallization robot. Diffraction-quality crystals were obtained from 0.025% dichloromethane/12% PEG 20,000/0.1M imidazole, pH7.0.

The crystals formed in the hexagonal space group $P3_112$, with cell constants a = 125.31, b = 125.31, c = 131.40 Å, and with one monomer per asymmetric unit (Figs. 11 and 12). A complete dataset, to 2.9 Å resolution, was collected at 100 K on a single crystal at the European Synchrotron Radiation Facility (ESRF) beam line, ID23-2 (wavelength 0.8726 Å). The diffraction images were indexed and integrated using the HKL2000 program (Otwinowski & Minor, 1997), and the integrated reflections were scaled using the SCALEPACK program (Otwinowski & Minor, 1997). The structure factor amplitudes were calculated using TRUNCATE from the CCP4 program suite (French & Wilson, 1978). The structure was solved by molecular replacement with the program PHASER (Murshudov et al, 1997), using as the starting model the refined structure of rhAChE complexed with fasciculin (Kryger et al, 2000) (PDB code 1B41). All steps of atomic refinement utilized the CCP4/REFMAC5 program (Murshudov et al, 1997). The model was built into $2mF_{obs} – DF_{calc}$, and $mF_{obs} - DF_{calc}$ maps using the COOT program (Emsley & Cowtan, 2004). Refinement weights were optimized. The hAChE construct contains 543 amino acid residues. The final model includes residues Glu5-Ser541 (Fig. 12). The R_{free} value is 23.74% (based on the 5% of reflections not used in the refinement), and the R_{work} value is 19.31% for all data to 2.9 Å. The hAChE model was evaluated with the PROCHECK program (Laskowski et al, 1993). The coordinates and structures factors have been deposited to the PDB and have been assigned PDB Code 4PQE.

The 4PQE crystal structure of rhAChE diffracts to 2.9 Å resolution, in the space group *$P3_112$*, and with one monomer in the asymmetric unit. The structure of rhAChE was previously solved at 3.2 Å resolution in space group *$P6_1$* (PDB code 3LII) (Dvir et al, 2010), and at 2.16 Å resolution in space group *$P3_121$* (PDB code 4EY4) (Cheung et al, 2012). In both cases the crystal structure revealed a dimer in the asymmetric unit. The structure of the hAChE monomer is characterized by an α/β-hydrolase fold that consists of a 12-stranded central mixed β-sheet surrounded by 14 α-helices. Three intramolecular disulfide bridges, Cys68-Cys95, Cys256-Cys271 and Cys408-Cys528, are seen. The functional hAChE dimer was generated using the two-fold crystallographic symmetry present in the *$P3_112$* space group. The subunits associate with each other by tight packing of two antiparallel α-helices from each subunit, α13 and α18-α19, which form a 4-helix bundle. The 4-helix bundle is a widely occurring structural motif for protein–protein interaction that takes place at the dimer interface between two monomers, each of which contributes two helices (Kohn et al, 1997; Lin et al, 1995). In many cases these helices are close to each other within the polypeptide sequence, separated from each other by three loops (Kresse et al, 2001). The interfaces between the helices consist mostly of hydrophobic residues, while polar side chains on the exposed surfaces interact with the aqueous environment. The macrodipoles of the two α-helices of the individual polypeptides are in a stabilizing anti-parallel arrangement (Fuxreiter & Simon, 2002; Robinson & Sligar, 1993). The 4-helix bundle is arranged in an up-and-down α13, α18-α19, α13', α18'-α19' topology ↓↑↓↑. As a consequence, the active-site gorges of the two subunits face in opposite directions (Fig. 13). The 4-helix bundle is further stabilized by the interactions between conserved residues in the loop connecting α13 to α14 and residues from helices α18'-α19', and, conversely, between residues in the loop connecting α13' to α14' and residues from helices α18-α19. Non-polar residues point into the core of the bundle, while polar residues are oriented towards the aqueous solvent.

It is well established that the overall folds observed for the same protein crystallized in different space groups are usually very similar. However, local details of the structure may differ, especially the conformations of loops and surface side-chains. Side-chain conformations are determined by a combination of intramolecular and intermolecular forces. Thus, differences in surface side-chain conformations from space group to space group are due to different crystal packing constraints and/or variations in the ionic strength of the precipitant.

We performed a structure-based sequence comparison of the crystal structure of hAChE solved in the new space group $P3_112$ with those of other AChEs (Table 1). hAChE crystal structures represent three different space groups, $P3_112$, $P6_1$ and $P3_121$ (PDB codes: 4PQE, 3LII and 4EY4, respectively) (Cheung et al, 2012; Dvir et al, 2010); mouse AChE (mAChE) has been crystallized in the $P2_12_12_1$ space group (PDB codes: 1MAA and 1J06, each with different cell dimensions) (Bourne et al, 1999b; Bourne et al, 2003); *Torpedo californica* AChE (*Tc*AChE) in three different space groups, $P2_12_12_1$, *C121* and $P3_121$, (PDB codes: 1W75, 2J3D and 1EA5, respectively) (Dvir et al, 2002; Greenblatt et al, 2004; Harel et al, 2006); *Drosophila melanogaster* AChE (*Dm*AChE) in space group $P4_32_12$ (PDB code 1QO9) (Harel et al, 2000), and *Electrophorus electricus* AChE (*Ee*AChE) in space groups *C121* and *F222.* Two *F222* crystal form were described, with different unit cell parameters (PDB codes: 1C2O, 1C2B and 1EEA respectively; Table 1). All three *Ee*AChE structures were solved to limited resolution (4.2, 4.5 and 4.5 Å, respectively) (Table 2), such that the *Ee*AChE residues could not be assigned unambiguously. Indeed, the coordinates deposited in the PDB are of the mAChE sequence for 1C2O and 1C2B (Bourne et al, 1999a; Bourne et al, 1999b; Bourne et al, 2003), and of the *Tc*AChE sequence for 1EEA (Raves et al, 1998; Raves et al, 1997). The *Ee*AChE structures were, therefore, omitted from this comparative structural analysis.

Thus, we compared the orientations of 9 crystal structures of AChE from four different species in six different crystal forms. In some of these structures the asymmetric unit contains one or two homodimers, while in others it contains a single monomer, and the dimer is formed by a crystallographic 2-fold axis. The various structures contain 4-16 protein molecules in the crystallographic unit cell, such that there are 2-4 different monomer-monomer contacts in each unit cell of the different crystal forms. Our analysis revealed that all the crystal forms contain at least one physiological AChE dimer in the unit cell, with tight packing of two helices from each monomer producing the canonical cholinesterase 4-helix bundle. However, the crystal contacts between any other two adjacent subunits in the crystallographic unit cell significantly differ from one pair to the other, generating different non-physiological non-specific dimeric associations. For example, mAChE crystallized in $P2_12_12_1$ space group (PDB code 1J06), with 8 subunits in the unit cell (labeled A-H) with one physiological dimer AB, and three dimers - CD, EF and GH - related by crystallographic symmetry (Fig. 14) (Bourne et al, 1999b). Several of the residues in the 4-helix bundle, including Glu376, Thr383, Asp384, Trp385, Gln508, Gln527, Phe535 and Lys538 (hAChE numbering), are strictly conserved in all three vertebrate species, but not in *Dm*AChE (Fig. 15). There are three distinct non-physiological dimers - AD, AG and AH in mAChE (Bourne et al, 1999b) (Fig. 14). Analysis of the contacts between these non-physiological dimers revealed that they involve non-conserved residues, and that the residues are different in each of the three pairs, giving rise to different non-specific dimer association. They also do not include any residues that are involved in forming the 4-helix bundle. These observations apply to all the different space groups in the various species. Furthermore, the crystal contacts are much looser, *i.e.*, involve fewer interactions and fewer residues, in each of the non-specific dimers of the AChE structures, relative to the very tight association in the physiological dimers. Thus, for human and mouse AChEs, the physiological dimer appears to be the one homologous to the one in *Tc*AChE. Both the conservation of residues and the tighter association support the notion that the dimers formed via the 4-helix bundle are, in fact, the physiological dimers (Table 1).

We evaluated the effect of crystal contacts on accessibility to the entrance of the active-site gorge. For each of the two subunits of the physiological dimer, we analyzed residues involved in close

contacts, *i.e.*, up to 3.5 Å, with residues from neighboring subunits in the unit cell of the various crystal structures. In 5 of the 9 AChE crystal structures the crystal contacts with residues from symmetry-related subunits are located far from the entrance to the gorge. In four crystal forms, hAChE (PDB 4PQE), mAChE (PDB code 1MAA), *Tc*AChE (PDB code 1EA5) and hAChE (PDB code 3LII), these residues are near the entrance to the gorge of one of the two subunits forming the physiological AChE dimer. *Tc*AChE crystallized in P3121 space group (PDB code 1EA5) has 6 subunits in the unit cell (A-F), forming five subunit-subunit contacts (Fig. 16a). The two monomers that form the physiological dimer (AE), also form non-physiological dimers (AB, AC, ED and EF). The crystal-contacts between subunits A (part of the physiological AE dimer) and C, and between D and E (part of the physiological AE dimer) involve residues situated at the peripheral anionic site (PAS) that might, therefore, block accessibility to the active site (Fig. 16b). In the hAChE (PDB codes: 4PQE and 3LII) and mAChE (PDB code: 1MAA) structures some of the residues involved in crystal contacts are also close to the entrance of the gorge. As a result, accessibility to the entrance of the gorge of the neighboring subunit is partially restricted.

To further evaluate the effect of neighboring copies of the catalytic subunit in the various space groups we compared the electrostatic characteristics of the AChE dimers. The electrostatic potentials for the AChE monomer structures from various species have a negative external surface potential in the area around the entrance to the active-site gorge, that becomes more negative as one approaches the rim of the gorge (Felder et al, 1997). The potential becomes increasingly more negative along the central axis running down the gorge, and is largest at the base of the gorge, near the active site. The active-site gorge of the two subunits in the physiological dimer face in opposite directions, consequently the negative external surface potentials of each monomer point in opposite directions, *i.e.*, the dipoles are related by the same 2-fold axis as the two monomers.

Since the AChE monomer displays an unusually large electrostatic dipole moment (Porschke et al, 1996), we were curious as to how it might affect the crystallographic packing and the accessibility of the entrance to the active-site gorge in the crystal. We calculated the electrostatic surface potentials of the individual AChE monomers in the different crystal forms. In *Tc*AChE the negative potential surface at the peripheral site of subunit E of the AE physiological dimer (black arrow in Fig. 16d) is adjacent to the positive potential surface of the symmetry-related subunit D (black arrow in Fig. 16c). Thus it is clear that the non-physiological dimer association between the ED subunits involves electrostatic complementarity.

The structures of the catalytic subunit monomers of the AChEs of the three vertebrate species, *h*AChE, mAChE and *Tc*AChE, show a very high degree of evolutionary conservation, displaying very limited conformational differences. *Dm*AChE displays somewhat larger deviations. The structural alignment of all the AChEs display root mean square deviations (RMSDs) that are consistent with the evolutionary distances and sequence differences between them (Fig. 15). Thus, hAChE shares ~88% sequence identity with mAChE, with an RMSD of 0.6-0.7 Å; 57% identity and an RMSD of 0.9 Å with *Tc*AChE, and an RMSD of 1.5 Å and 35.9% identity with *Dm*AChE (Table 1). The most pronounced differences between hAChE and *Dm*AChE are seen in the loop connecting α13 to α14 of the 4-helix bundle (residues Tyr382-Asp390 in hAChE and Tyr416-Gly424 in *DmA*ChE), with deviations ranging from 2.3 Å to 8.0 Å. In the vertebrate AChEs this loop points towards and close to its neighboring subunit, whereas in the *Dm*AChE structure it is positioned further away from the helix bundle axis, leading to misalignment in the dimer structure as compared to hAChE (Fig. 17).

3. KEY RESARCH ACCOMPLISHMENTS

- Cloning and expression of the WT rhAChE monomer in HEK293 cells
- Purification and characterization of rhAChE
- Deglycosylation of the expressed rhAChE
- Crystallization of the expressed rhAChE
- X-ray data collection of rhAChE
- 3D structure determination of rhAChE
- Cloning and expression of rhAChE$_T$ in HEK293 cells
- Affinity purification of rhAChE$_T$
- Characterization of rhAChE$_T$
- Deglycosylation of rhAChE$_T$
- Tetramerization of DG-rhAChE$_T$

4. REPORTABLE OUTCOMES

Presentations:

- 14th ISCM in Hangzhou, China (May 2013) by Joel L. Sussman, title: “Acetylcholinesterase: From 3D Structure to Dynamics
- ICSG2013-SLS, Sapporo, Japan (Jul 2013) by Joel L. Sussman, title: “Cholinesterases: From 3D Structure to Dynamics”
- Invited Lecture at Regional Centre for Biotechnology, Gurgaon, India (Sep, 2013) by Joel L. Sussman, title: “Molecular Basis of How Nerve Agents through anti-Alzheimer Drugs Function: 3D Structure of Acetylcholinesterase”
- Regional Centre for Biotechnology, Gurgaon, India (Sep 2013) - Keynote Lecture by Joel L. Sussman, title: “Molecular Basis of How Nerve Agents through anti-Alzheimer Drugs Function: 3D Structure of Acetylcholinesterase
- Florida International University (FIU), Miami, FL (Dec 2013) - Invited Lecture by Joel L. Sussman, title: “Molecular Basis of anti-Alzheimer Drugs & Nerve Agents: 3D Structure of Acetylcholinesterase”
- Technion, Haifa, ISRAEL (Dec 2013) - Invited Lecture by Joel L. Sussman, title: “Molecular Basis of How Nerve Agents & anti-Alzheimer Drugs Function: 3D Structure of Acetylcholinesterase”

5. CONCLUSIONS

The human acetylcholinesterase (hAChE) gene was cloned into the pHLsec expression vector. It was expressed in large-scale utilizing 64 large plates (15 cm) of adherent 293T cells. The recombinant human AChE (rhAChE) was purified by affinity chromatography, and deglycosylated with PNGase F, resulting in crystals diffracting to 2.9 Å that permitted 3D structure determination and refinement. The data were deposited in the PDB (IDcode 4PQE).

The major outcome of the structural information provided by this research is the potential to develop one or more products based on a combination of hAChE with one or more oxime reactivators that will act synergistically to perform pseudocatalytic hydrolysis of all V- or G-agents. It is envisaged that these products will be utilized either prophylactically or in a post-treatment context. Embedded in a suitable medical doctrine they will expedite the recovery and return to full battle capability of the war-fighter. The approach adopted should substantially reduce the amount of the enzyme protein required to afford protection, thus making its availability more economically feasible. The civilian population may be exposed to OP nerve agents both in military and terrorist scenarios. However, civilians are also exposed to and susceptible to a variety of commercially available OP-based and carbamate-based insecticides in both a domestic and an agricultural context. The same considerations outlined above may thus also be pertinent to treatment of civilians. The issue of reduced protein dosage, resulting in enhanced cost effectiveness, will be particularly relevant to the civilian arena, where

treatment may be administered to a large and heterogeneous population in terms of age and medical background.

Future Plans

- To increase the yield of the recombinant protein by testing the expression in transiently transfected 293E cells, and to grow large quantities of culture in suspension.
- To simplify the protocol of purification with and without de-glycosylation.
- To purify and characterize the rhAChEG4 construct.
- Initiate experiments on soaking OP surrogates into native crystals of rhAChEmD already obtained, and subsequently into rhAChEG4 crystals, which likely will correspond to the physiological tetramer.
- To determine the kinetics of interaction of the nerve agent surrogates with the expressed constructs, as well as the rates of aging of the inactive OP conjugates thus obtained, and of their reactivation by suitable oximes.
- Initiation of crystallization screens of G-agent/rhAChE conjugates obtained by interaction of the enzyme with G-agent surrogates.

6. REFERENCES

Bourne Y, Grassi J, Bougis PE, Marchot P (1999a) Conformational flexibility of the acetylcholinesterase tetramer suggested by X-ray crystallography. *J Biol Chem* **274:** 30370-30376

Bourne Y, Taylor P, Bougis PE, Marchot P (1999b) Crystal structure of mouse acetylcholinesterase. A peripheral site-occluding loop in a tetrameric assembly. *J Biol Chem* **274:** 2963-2970

Bourne Y, Taylor P, Radic Z, Marchot P (2003) Structural insights into ligand interactions at the acetylcholinesterase peripheral anionic site. *EMBO J* **22:** 1-12

Cheung J, Rudolph MJ, Burshteyn F, Cassidy MS, Gary EN, Love J, Franklin MC, Height JJ (2012) Structures of human acetylcholinesterase in complex with pharmacologically important ligands. *J Med Chem* **55:** 10282-10286

Dvir H, Harel M, Bon S, Liu W-Q, Vidal M, Garbay C, Sussman JL, Massoulié J, Silman I (2004) The synaptic acetylcholinesterase tetramer assembles around a polyproline-II helix. *EMBO J* **23:** 4394-4405

Dvir H, Jiang HL, Wong DM, Harel M, Chetrit M, He XC, Tang XC, Silman I, Bai DL, Sussman JL (2002) X-ray structures of *Torpedo californica* acetylcholinesterase complexed with (+)-Huperzine A and (-)-Huperzine B: Structural evidence for an active site rearrangement. *Biochemistry* **41:** 10810-10818

Dvir H, Silman I, Harel M, Rosenberry TL, Sussman JL (2010) Acetylcholinesterase: From 3D structure to function *Chem Biol Interact* **187:** 10-22

Ellman GL, Courtney KD, Andres V, Featherstone RM (1961) A new and rapid colorimetric determination of acetylcholinesterase activity. *Biochem Pharmacol* **7:** 88-95

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60:** 2126-2132

Felder CE, Botti SA, Lifson S, Silman I, Sussman JL (1997) External and internal electrostatic potentials of cholinesterase models. *J Molec Graphics & Modelling* **15:** 318-327

French S, Wilson K (1978) On the treatment of negative intensity observations. *Acta Crystallographica Section A* **34:** 517-525

Fuxreiter M, Simon I (2002) Role of stabilization centers in 4 helix bundle proteins. *Proteins* **48:** 320-326

Giles K, Ben-Yohanan R, Velan B, Shafferman A, Sussman JL, Silman I (1998) Assembly of acetylcholinesterase subunits *in vitro*. In *Structure and Function of Cholinesterases and Related Proteins*, Doctor BP, Quinn DM, Rotundo RL, Taylor P (eds), pp 442-443. New York: Plenum

Greenblatt HM, Guillou C, Guénard D, Argaman A, Botti S, Badet B, Thal C, Silman I, Sussman JL (2004) The complex of a bivalent derivative of Galanthamine with *Torpedo* acetylcholinesterase displays drastic deformation of the active-site gorge: Implications for structure-based drug design. *J Am Chem Soc* **126:** 15405-15411

Grueninger-Leitch F, D'Arcy A, D'Arcy B, Chene C (1996) Deglycosylation of proteins for crystallization using recombinant fusion protein glycosidases. *Protein Sci* **5:** 2617-2622

Harel M, Kryger G, Rosenberry TL, Mallender WD, Lewis T, Fletcher RJ, Guss JM, Silman I, Sussman JL (2000) Three-dimensional structures of *Drosophila melanogaster* acetylcholinesterase and of its complexes with two potent inhibitors. *Protein Sci* **9:** 1063-1072

Harel M, Silman I, Sussman JL. (2006) Native monoclinic form o Torpedo acetylcholinesterase.

Kohn WD, Mant CT, Hodges RS (1997) Alpha-helical protein assembly motifs. *J Biol Chem* **272:** 2583-2568

Kresse HP, Czubayko M, Nyakatura G, Vriend G, Sander C, Bloecker H (2001) Four-helix bundle topology re-engineered: monomeric Rop protein variants with different loop arrangements. *Protein Eng* **14:** 897-901

Kronman C, Chitlaru T, Elhanany E, Velan B, Shafferman A (2000) Hierarchy of post-translational modifications involved in the circulatory longevity of glycoproteins. Demonstration of concerted contributions of glycan sialylation and subunit assembly to the pharmacokinetic behavior of bovine acetylcholinesterase. *J Biol Chem* **275:** 29488-29502

Kryger G, Harel M, Giles K, Toker L, Velan B, Lazar A, Kronman C, Barak D, Ariel N, Shafferman A, Silman I, Sussman JL (2000) Structures of recombinant native and E202Q mutant human acetylcholinesterase complexed with the snake-venom toxin fasciculin-II. *Acta Crystallogr D Biol Crystallogr* **56:** 1385-1394

Laskowski RA, MacArthur MW, Moss D, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* **26:** 283-291

Lin SL, Tsai CJ, Nussinov R (1995) A study of four-helix bundles: investigating protein folding via similar architectural motifs in protein cores and in subunit interfaces. *J Mol Biol* **248:** 151-161

Murshudov G, Vagin A, Dodson E (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53:** 240-255

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276:** 307-326

Porschke D, Créminon C, Cousin X, Bon C, Sussman JL, Silman I (1996) Electrooptical measurements demonstrate a large permanent dipole moment associated with acetylcholinesterase. *Biophys J* **70:** 1603-1608

Raves M, Giles K, Schrag JD, Schmid MF, Phillips GN, Chiu W, Howard AJ, Silman I, Sussman JL (1998) Quaternary structure of tetrameric acetylcholinesterase. In *Structure and Function of Cholinesterases and Related Proteins*, Doctor BP, Taylor P, Quinn DM, Rotundo RL, Gentry MK (eds), pp 351-356. New York: Plenum

Raves ML, Harel M, Pang Y-P, Silman I, Kozikowski AP, Sussman JL (1997) 3D structure of acetylcholinesterase complexed with the nootropic alkaloid, (-)-huperzine A. *Nat Struct Biol* **4:** 57-63

Robinson CR, Sligar SG (1993) Electrostatic stabilization in four-helix bundle proteins. *Protein Sci* **2:** 826-837

Sussman JL, Harel M, Frolow F, Varon L, Toker L, Futerman AH, Silman I (1988) Purification and crystallization of a dimeric form of acetylcholinesterase from *Torpedo californica* subsequent to solubilization with phosphatidylinositol-specific phospholipase C. *J Mol Biol* **203:** 821-823

Velan B, Kronman C, Grosfeld H, Leitner M, Gozes Y, Flashner Y, Sery T, Cohen S, Ben-Aziz R, Seidman S, Shafferman A, Soreq H (1991) Recombinant human acetylcholinesterease is secreted from transiently transfected 293 cells as a soluble globular enzyme. *Cell Mol Neurobiol* **11:** 143-156

7. APPENDICES

None.

8. SUPPORTING DATA

Tables:

Table 1: Crystal structures and crystal contacts for 3D AChE structures

Species	Space Group	Crystal Form	PDB	Seq ID (%)	# dimer contacts	Number of crystallographic contacts	AU	UC
human	$P3_111$	trigonal	4PQE	100	26	14	1	6
human	$P6_1$	hexagonal	3LII	100	26	11	2	12
human	$P3_121$	trigonal	4EY4	100	31	3-16	2	12
mouse	$P2_12_12_1$	orthorhombic	1J06	88.6	30	6-20	2	8
mouse	$P2_12_12_1$	orthorhombic	1MAA	88.6	26	10-22	4	16
Torpedo	$P2_12_12_1$	orthorhombic	1W75	57.5	11	2-8	2	8
Torpedo	C121	monoclinic	2J3D	57.5	23	19	1	4
Torpedo	$P3_121$	trigonal	1EA5	57.5	21	19	1	6
Electrophorus[1]	C121	monoclinic	1C2O	88.5	--	--	4	16
Electrophorus[1]	F222	orthorhombic	1C2B	88.5	--	--	1	16
Electrophorus[2]	F222	orthorhombic	1EEA	88.5	--	--	-	--
Drosophila	$P4_32_12$	tetragonal	1QO9	35.9	25	2-15	1	8

[1] In the 1C2O and 1C2 PDB structures the sequence of mAChE was used due to the low resolution of the X-ray data; thus it does not seem useful to generate a list of crystallographic contacts, etc.

[2] In the 1EEA PDB structure the sequence of *Tc*AChE was used due to the low resolution of the X-ray data; thus it does not seem useful to generate a list of crystallographic contacts, etc.

Table 2: Crystallization and crystallographic data for crystal forms of AChEs (continued)

Species (PDB ID Code)	**Space Group**	**Cell Dimensions**	**RMSD to new Å**	**Res Å**	**Seq identity to human (%)**	**Crystallization conditions**
Human (4PQE)	$P3_111$	a=b= 125.31 c=131.39 γ=120		2.90	100	0.1 M imidazole, pH=7 12% PEG 20,000 0.5% ethyl acetate
Human (3LII)	$P6_1$	a=b= 210.90 c=115.27 γ=120	0.525	3.20	100	1.3–1.5 M $(Li)_2SO_4$ 0.1 M HEPES, pH=6.8
Human (4EY4)	$P3_121$	a=b= 106.18 c=324.39 γ=120	0.502	2.16	100	15-21% PEG 3,350 0.17-0.21 M KNO_3
Mouse (1J06)	$P2_12_12_1$	a=79.19 b=111.97 c=226.83 α=β= γ=90	0.635	2.35	88.6	25-32% PEG 600 25-100 mM HEPES or sodium acetate, pH 6.5–8.0
Mouse (1MAA)	$P2_12_12_1$	a=136.55 b=173.13 c=224.25 α=β=γ=90	0.632	2.90	88.6	1.7 M $NaKPO_4$, pH=7.0 10 mM $CaCl_2$
Torpedo (1W75)	$P2_12_12_1$	a=91.50 b=106.58 c=150.34 α=β=γ=90	0.632	2.40	57.5	40% PEG200 0.5M MES pH=5.8
Torpedo (2J3D)	C121	a=128.74 b=105.49 c=70.60 β= 106.62	0.947	2.60	57.5	30% PEG 400 0.1M Tris pH=8.5 0.2M $MgCl_2$ PEG 550MME
Torpedo (1EA5)	$P3_121$	a=b= 111.57 c=137.59 γ=120	0.947	1.80	57.5	38% PEG 200 0.1 M MES, pH=8.0
Electrophorus (1C2O)	C121	a=211.16 b=129.75 c=195.42 β= 103.20	0.691	4.20	88.5	1.1–1.5 M $NaKPO_4$ pH 8.0–9.5
Electrophorus (1C2B)	F222	a=117.98 b=215.87 c=229.41 α=β=γ=90	0.691	4.50	88.5	1.4 M $(NH_4)_2SO_4$, pH 5.5–6.0
Electrophorus (1EEA)	F222	a=140.86 b=201.46 c=237.77 α=β=γ=90	0.738	4.50	88.5	500 mM NaCl
Drosophila (1QO9)	$P4_32_12$	a=b= 94.34 c=159.00 α=β=γ=90	1.463	2.70	35.9	13% PEG MME 2,000 0.1 M $(NH_4)_2SO_4$ 0.03 M leucine 0.1 M sodium acetate, pH=4.6

Figures:

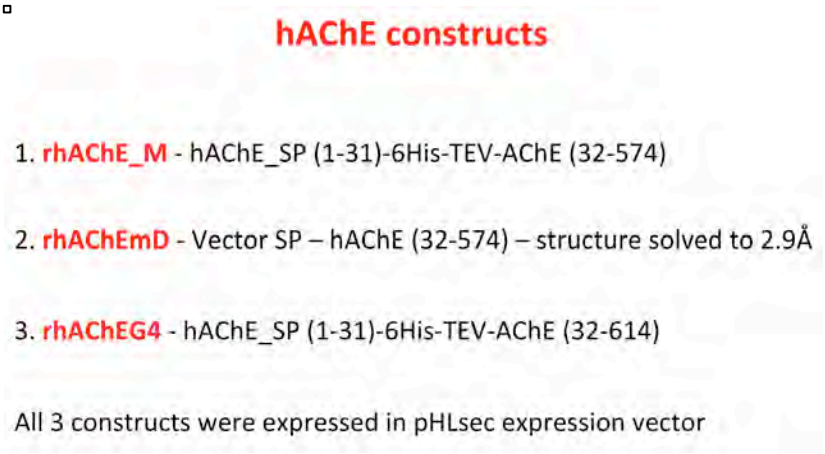

Fig. 1. hAChE constructs

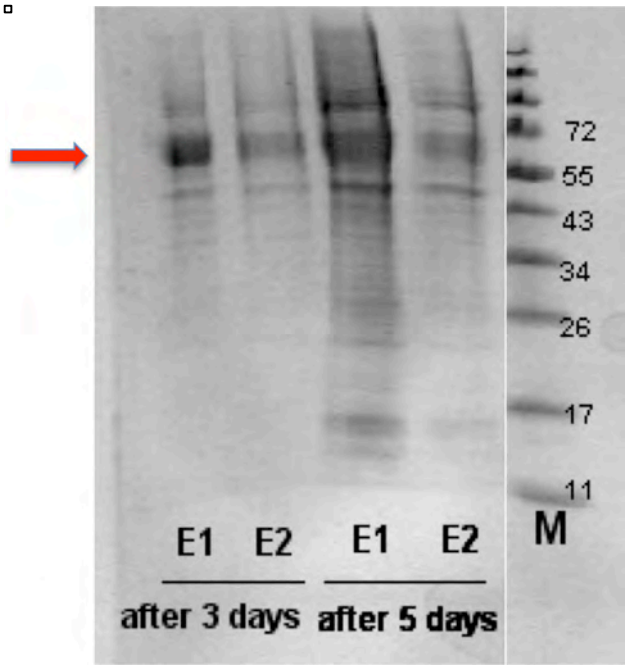

Fig. 2. Small-scale expression of rhAChE. Expression was tested in 293T cells, and the protein was secreted into the medium

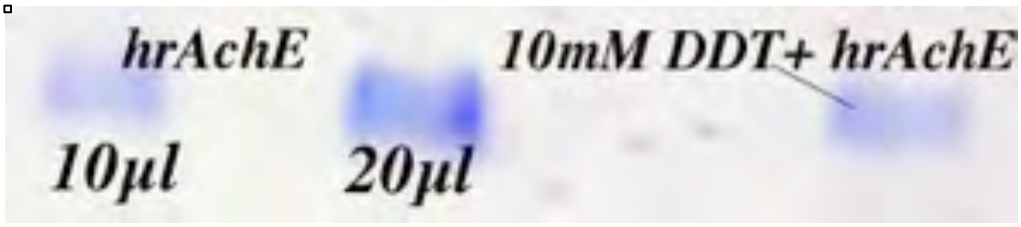

Fig. 3. Non-denaturing PAGE gel for purified rhAChE in the absence (left & center) and presence (right) of DTT.

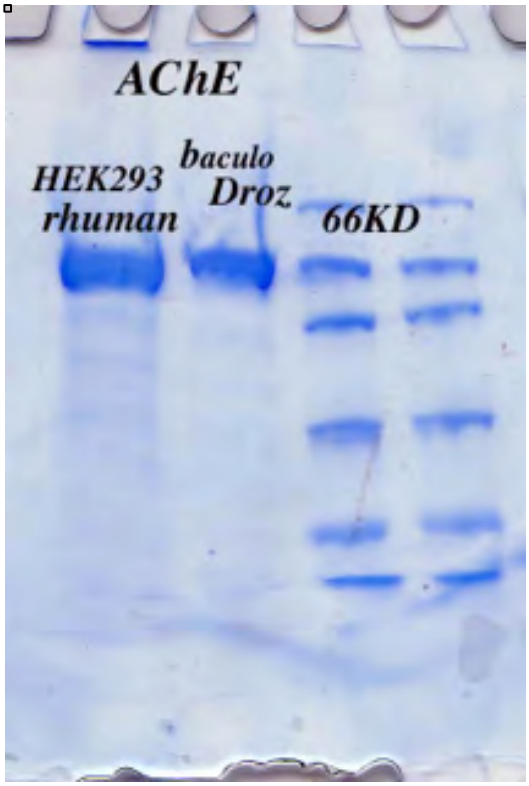

Fig. 4. SDS-PAGE on 7% acrylamide of rhAChE expressed in HEK293 (lane 1), of rhAChE expressed in baculovirus (lane 2), and of molecular weight markers (lanes 3 and 4).

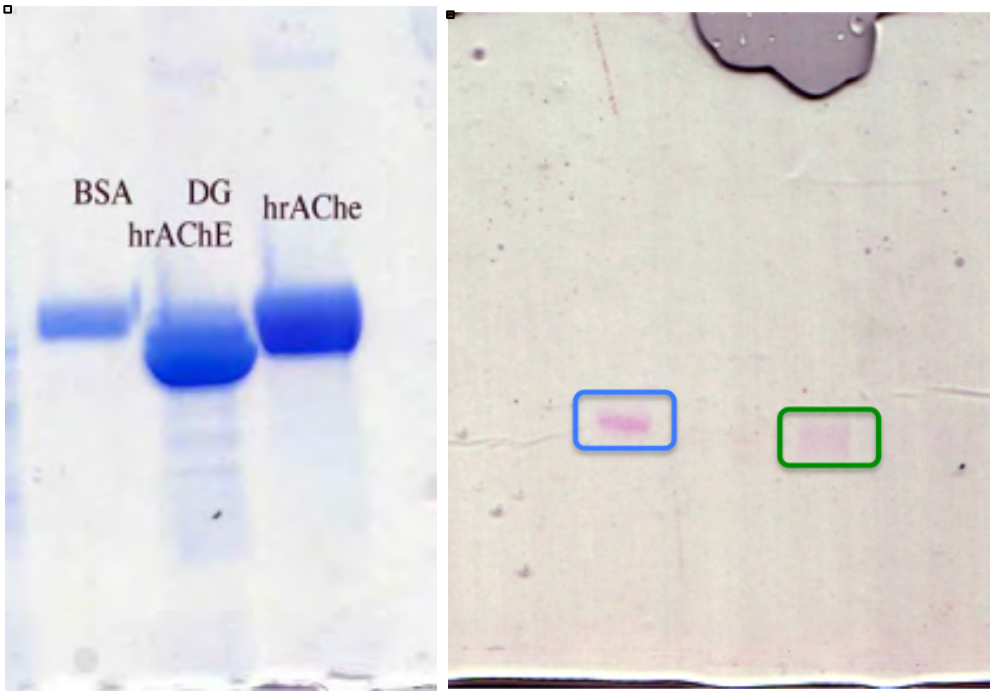

Fig. 5. SDS-Page under reducing conditions on 7% acrylamide of rhAChEmD before and after deglycosylation. Left panel: Lane 1 – BSA marker; Lane 2 – deglycosylated rhAChEmD; Lane 3 – non-deglycosylated rhAChEmD. Right panel: Lane 1 - non-deglycosylated rhAChEmD [marked with blue rectangle]; Lane 2 - deglycosylated rhAChEmD, marked with green rectangle, both revealed with GlycoStain).

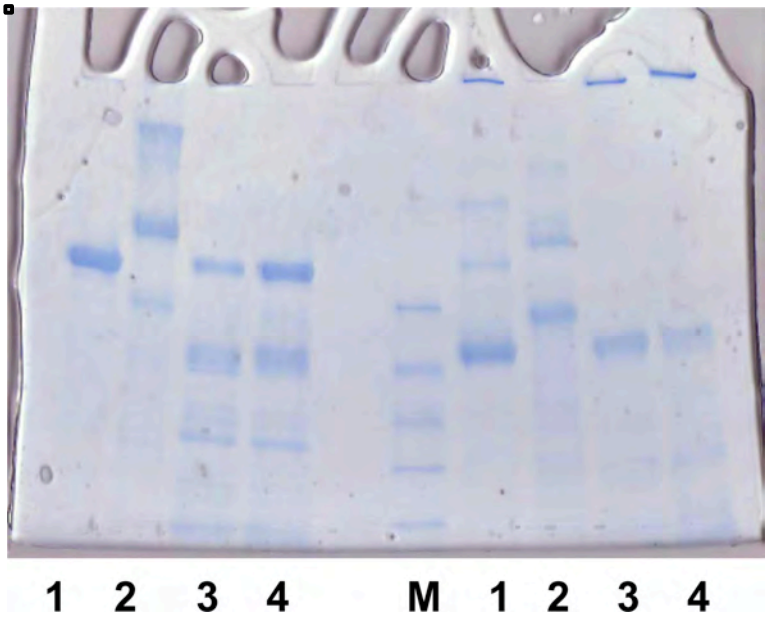

Fig. 6. SDS-PAGE of purified rhAChE$_T$. Electrophoresis was performed on 4-15% gradient gels under non-reducing (NR) and reducing (R) conditions, reduction being achieved with β-mercaptoethanol. Under both conditions the rhAChE$_T$ is in lanes 3 and 4. Purified *Tc*AChE (lane 1) and hBChE (lane 2) are shown for comparison. The middle lane (M) contains molecular weight markers. Staining was with Gelcode Blue Stain.

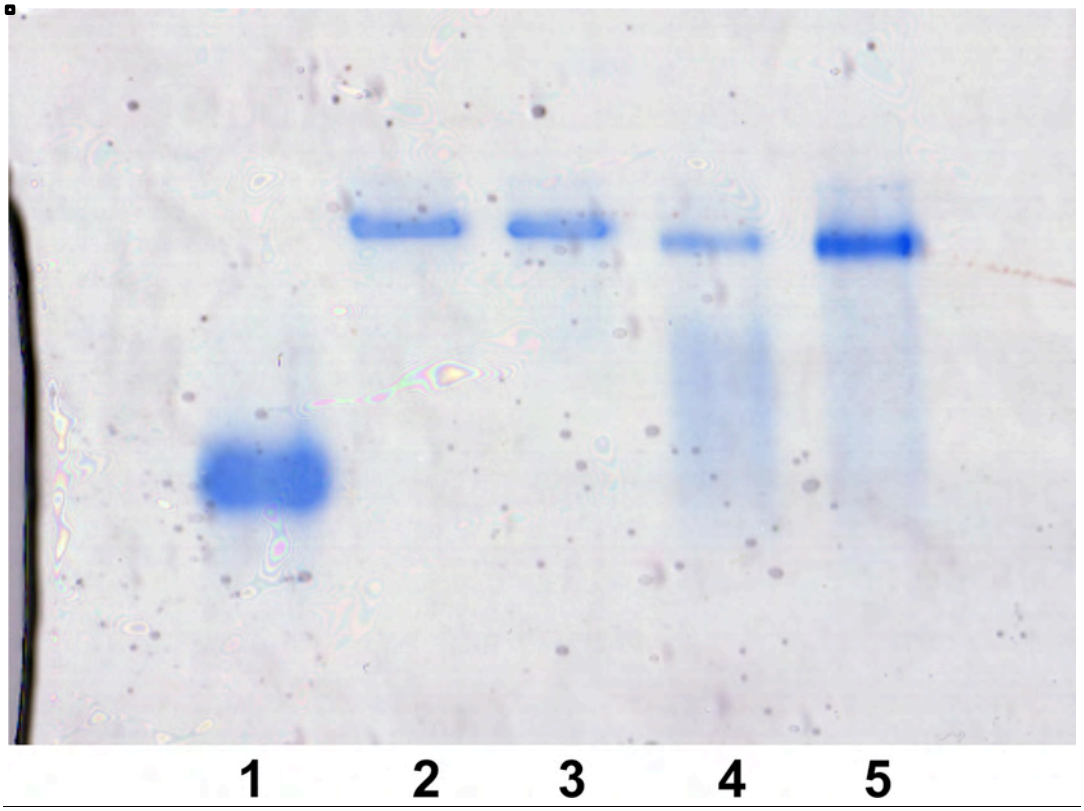

Fig. 7. Non-denaturing PAGE of purified rhAChE$_T$. Electrophoresis was performed on 7% gradient gels. Lane 1, *Tc*AChE; lanes 2 & 3, hBChE; lanes 4 & 5, rhAChE$_T$. Staining was with Gelcode Blue Stain.

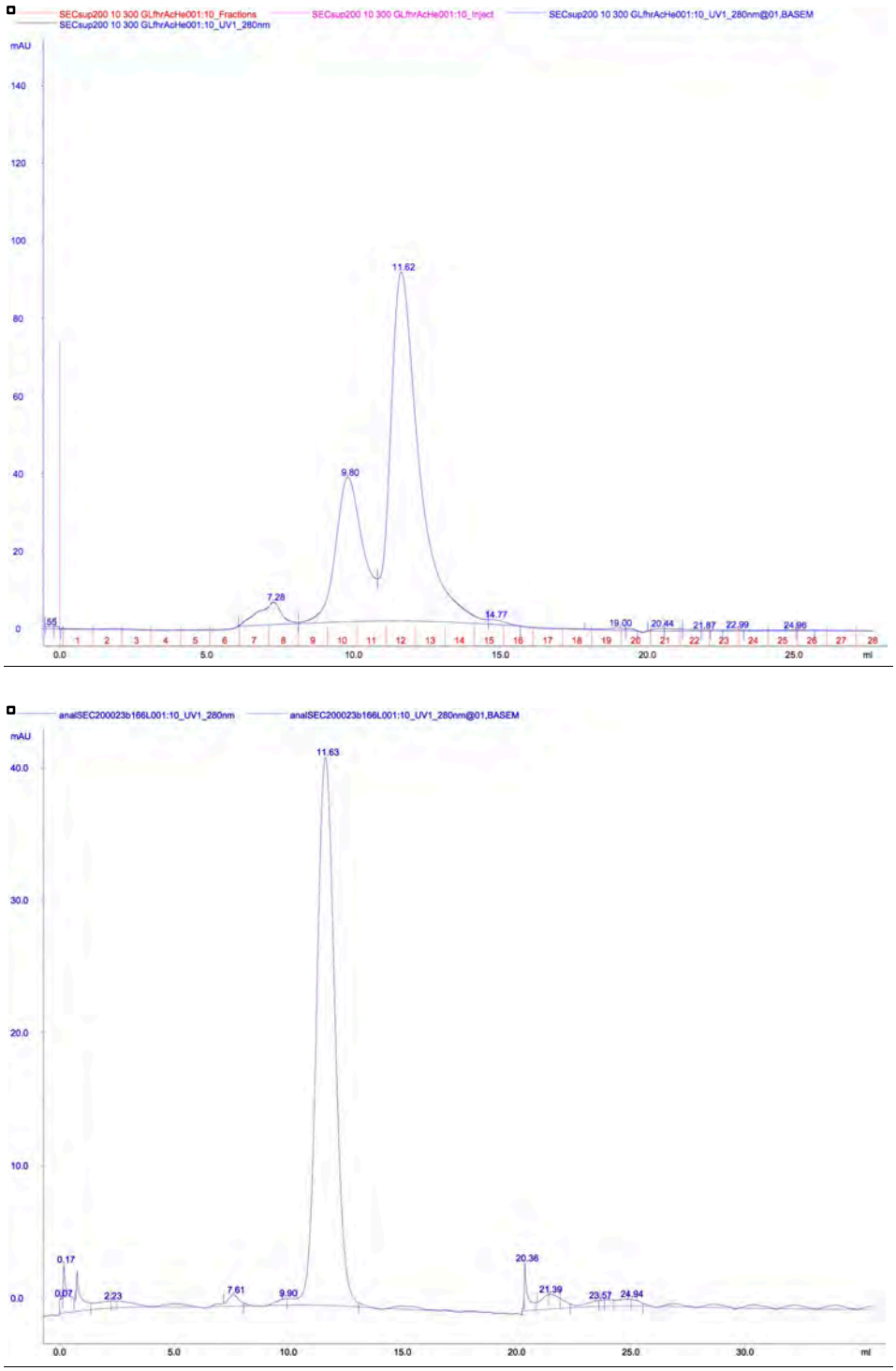

Fig. 8. SEC of affinity-purified AChE samples. SEC was performed on a Superdex 200 10/300 column, monitoring for protein at 280 nm. The buffer employed was 0.3 M NaCl/20 mM Tris, pH 8.0. Upper panel, rhAChE$_T$; lower panel, G$_2$ *Tc*AChE.

WAT: DTLDEAERQWKAEFHRWSSYMVHWKNQFDHYSKQDRCDL
PRAD: LLTPPPPPLFPPPFF

Fig. 9. Sequences of the human WAT and PRAD polypeptide sequences

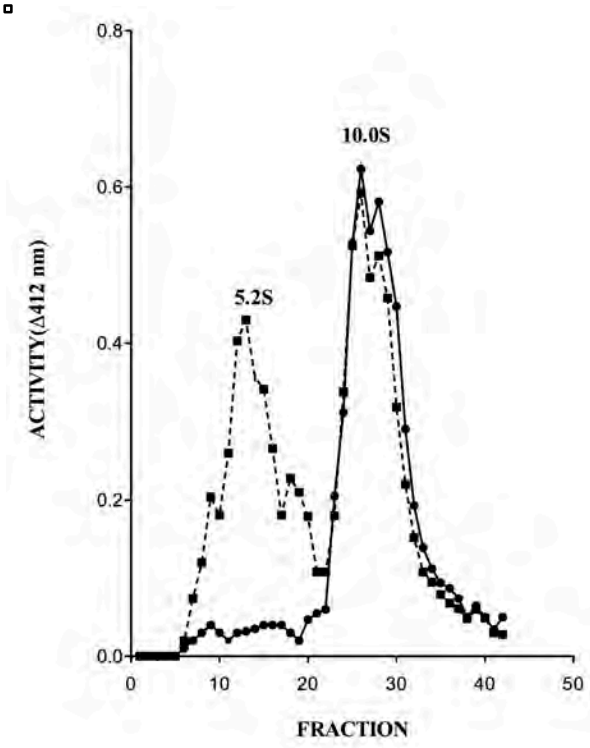

Fig. 10. Sucrose gradient centrifugation of DG-rhAChE$_T$. Centrifugation was on 5-20% sucrose gradients in 1% Triton X100/PBS, pH 7.0. ■−−■−−■, DG-rhAChE$_T$ in the absence of PRAD; ●-●-●, DG-rhAChE$_T$ preincubated with PRAD.

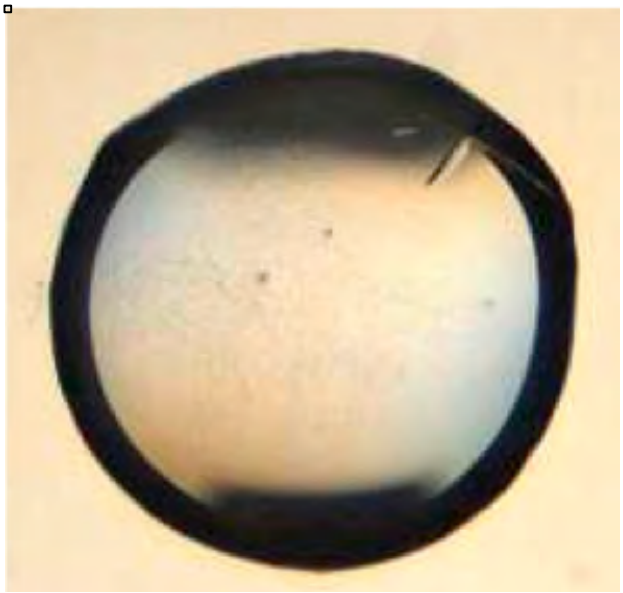

Fig. 11. Crystals of rhAChEmD, grown in sitting drops from a solution of 100 mM imidazole pH=7.0/12% polyethylene glycol 20,000, diffracted to 2.9Å resolution.

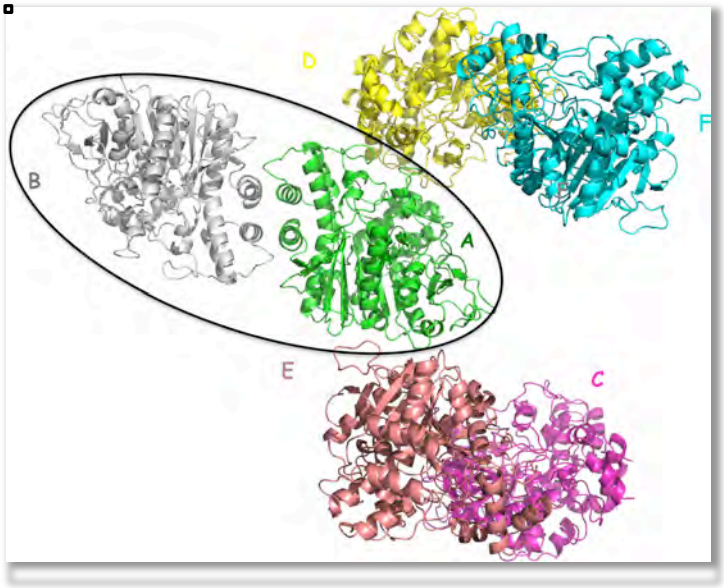

Fig. 12. Crystal packing of rhAChEmD solved to 2.9Å in space group $P3_112$

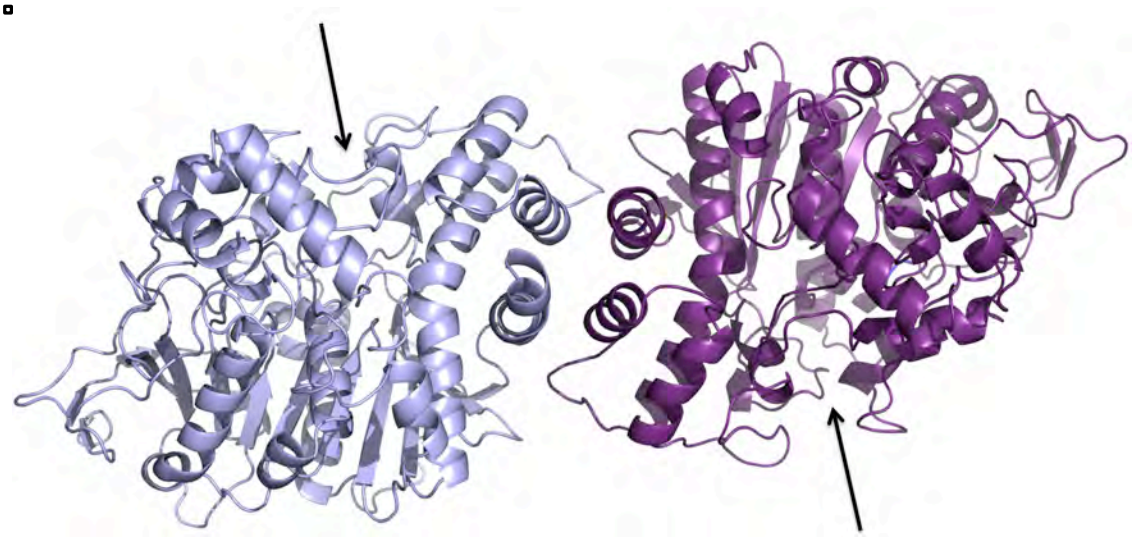

Fig. 13. Molecular dimer as seen in the rhAChE crystal structure. View is down the crystallographic 2-fold axis, showing the 4-helix bundle, with black arrows indicating the entrances to the active-site gorges of the two subunits facing in opposite directions

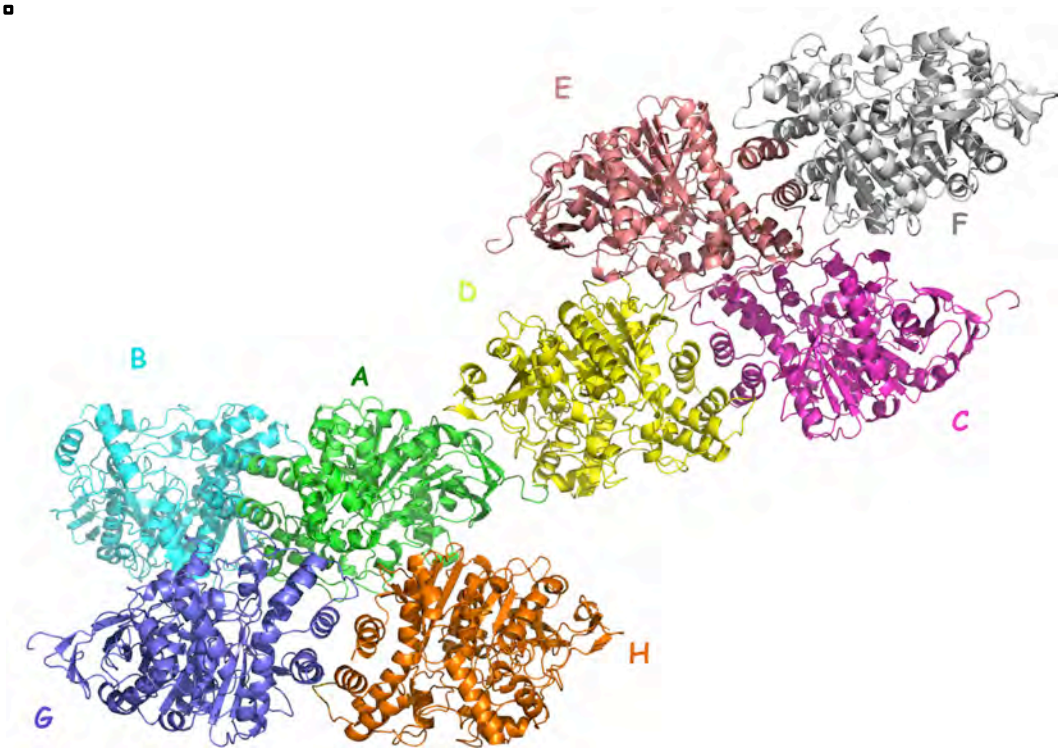

Fig. 14. Symmetry-related molecules in the $P2_12_12_1$ space group of mAChE (PDB1JO6)

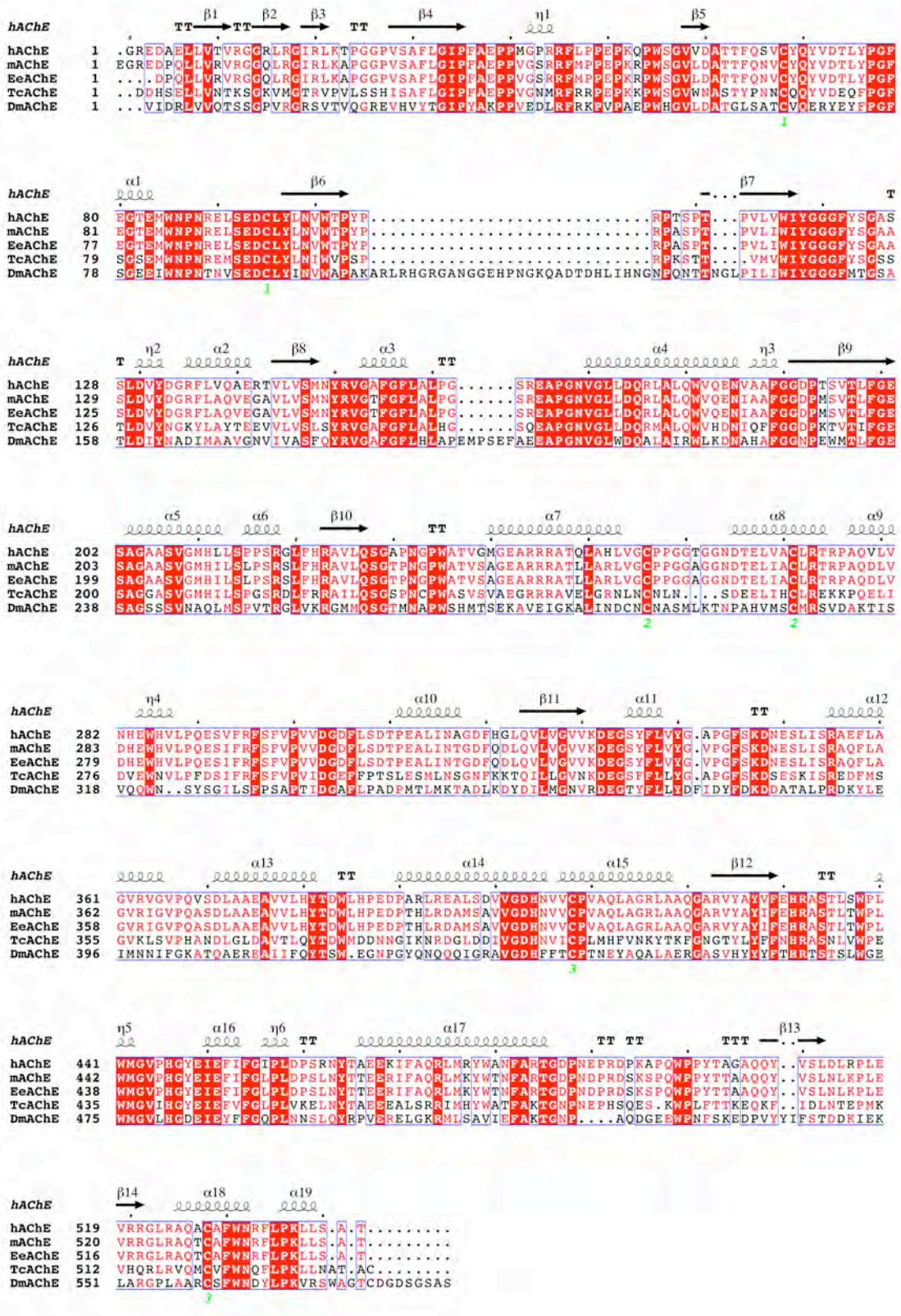

Fig. 15. Sequence alignment of AChEs whose 3D structures have been solved

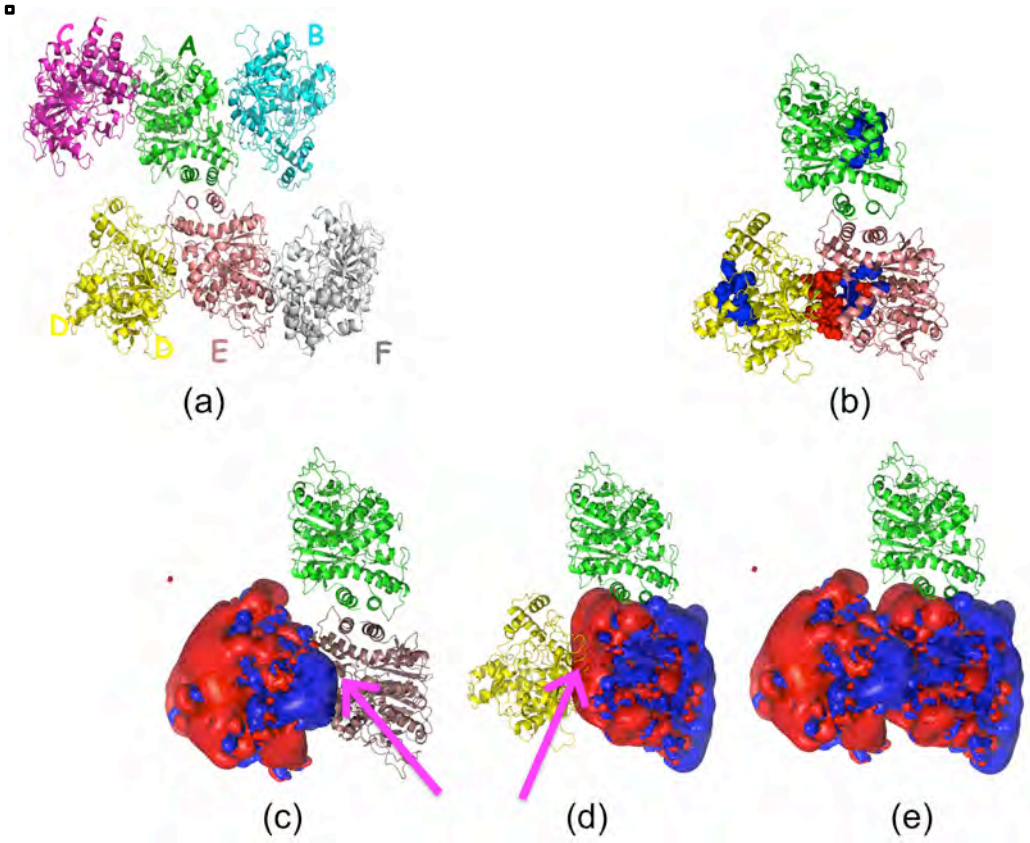

Fig. 16. Crystallographic packing and electrostatic interactions of *Tc*AChE (PDB Code 1EA5). (a) Crystallographic packing in space group $P3_121$; (b) representation in which the active-site residues are shown as blue balls, and the crystal contacts at the entrance to the gorge are shown in red; (c) monomer D (colored yellow in (b)), with its electrostatic surface colored red for negative potential and blue for positive potential, which forms a non-physiological contact with monomer E, with a pink arrow pointing towards the peripheral anionic site on monomer E; (d) monomer E (colored violet in (b)), with its electrostatic surface colored red for negative potential and blue for positive potential, which forms a non-physiological contact with monomer D, with a pink arrow pointing towards the peripheral anionic site on monomer E; (e) monomers D and E with their electrostatic surfaces colored red for negative potential and blue for positive potential.

Fig. 17. Alignment of rhAChE (cyan) and *Dm*AChE (orange), displaying the tight alignment of the β sheets (blue) and of the 4-helix bundle (purple), but larger divergence of the rest of the helices. The region with the most pronounces differences between hAChE and *Dm*AChE is in the loop connecting α13 to α14 of the 4-helix bundle (residues Tyr382-Asp390 in hAChE and Tyr416-Gly424 in *Dm*AChE), and is outlined in magenta.